

PCT

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/763,076

DATE: 03/02/2001
 TIME: 13:01:10

Input Set : A:\PPD50348 US SEQ LIST.txt
 Output Set: N:\CRF3\03022001\I763076.raw

p.5

3 <110> APPLICANT: Broekaert, Willem
 4 Francois, Isabelle
 5 Evans, Ian
 6 De Bolle, Miguel
 7 Ray, John
 9 <120> TITLE OF INVENTION: Genetic Method For The Expression Of Polyproteins In
 10 Plants
 12 <130> FILE REFERENCE: PPD50348/UST
 C--> 14 <140> CURRENT APPLICATION NUMBER: US/09/763,076
 C--> 15 <141> CURRENT FILING DATE: 2001-02-12
 17 <150> PRIOR APPLICATION NUMBER: GB 9818001.1
 18 <151> PRIOR FILING DATE: 1998-08-18
 20 <150> PRIOR APPLICATION NUMBER: GB 9826753.7
 21 <151> PRIOR FILING DATE: 1998-12-04
 23 <150> PRIOR APPLICATION NUMBER: PCT/GB99/02716
 24 <151> PRIOR FILING DATE: 1999-08-17
 26 <160> NUMBER OF SEQ ID NOS: 81
 28 <170> SOFTWARE: PatentIn Ver. 2.1
 30 <210> SEQ ID NO: 1
 31 <211> LENGTH: 446
 32 <212> TYPE: DNA
 33 <213> ORGANISM: Dahlia merckii
 35 <400> SEQUENCE: 1
 36 atggtgaatc ggtcggttgc gttctccgcg ttctgttctga tctttttcgt gctcgccatc 60
 37 tcaggttatc aaatcttttag ttcattttatt gaatatgata gtattttatat tcttttatgg 120
 38 ttttatgtgt tctgacaagt tgcaaatatt gagtagatat cgcattccgtt agtgaggaaac 180
 39 tatgcgagaa agctagcaag acatgggtcgg gaaactgtgg caatacggga cattgtgaca 240
 40 accaatgtaa atcatgggag ggtgcggccc atggagcgtg tcatgtgcgt aacgggaaac 300
 41 acatgtgttt ctgttacttc aattgtaaaa aagccgaaaa gcttgctcaa gacaaactta 360
 42 aagccgaaca actcgtctaa gacaaactta atgccccaaa gcttgaccgt gatgccaaga 420
 43 aagtgtttcc aaacgttgaa catccg 446
 47 <210> SEQ ID NO: 2
 48 <211> LENGTH: 118
 49 <212> TYPE: PRT
 50 <213> ORGANISM: Dahlia merckii
 52 <400> SEQUENCE: 2
 53 Met Val Asn Arg Ser Val Ala Phe Ser Ala Phe Val Leu Ile Leu Phe
 54 1 5 10 15
 56 Val Leu Ala Ile Ser Asp Ile Ala Ser Val Ser Gly Glu Leu Cys Glu
 57 20 25 30
 59 Lys Ala Ser Lys Thr Trp Ser Gly Asn Cys Gly Asn Thr Gly His Cys
 60 35 40 45
 62 Asp Asn Gln Cys Lys Ser Trp Glu Gly Ala Ala His Gly Ala Cys His
 63 50 55 60
 65 Val Arg Asn Gly Lys His Met Cys Phe Cys Tyr Phe Asn Cys Lys Lys
 66 65 70 75 80
 68 Ala Glu Lys Leu Ala Gln Asp Lys Leu Lys Ala Glu Gln Leu Ala Gln

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/763,076

DATE: 03/02/2001

TIME: 13:01:10

Input Set : A:\PPD50348 US SEQ LIST.txt

Output Set: N:\CRF3\03022001\I763076.raw

```

69          85          90          95
71 Asp Lys Leu Asn Ala Gln Lys Leu Asp Arg Asp Ala Lys Lys Val Val
72          100          105          110
74 Pro Asn Val Glu His Pro
75          115
79 <210> SEQ ID NO: 3
80 <211> LENGTH: 16
81 <212> TYPE: PRT
82 <213> ORGANISM: Artificial Sequence
84 <220> FEATURE:
85 <223> OTHER INFORMATION: Description of Artificial Sequence: Linker
86     propeptide
88 <400> SEQUENCE: 3
89 Ser Asn Ala Ala Asp Glu Val Ala Thr Pro Glu Asp Val Glu Pro Gly
90  1      5      10      15
94 <210> SEQ ID NO: 4
95 <211> LENGTH: 20
96 <212> TYPE: PRT
97 <213> ORGANISM: Artificial Sequence
99 <220> FEATURE:
100 <223> OTHER INFORMATION: Description of Artificial Sequence: Linker
101     propeptide
103 <400> SEQUENCE: 4
104 Lys Lys Ala Glu Lys Leu Ala Gln Asp Lys Leu Lys Ala Glu Gln Leu
105  1      5      10      15
107 Ile Gly Lys Arg
108          20
112 <210> SEQ ID NO: 5
113 <211> LENGTH: 40
114 <212> TYPE: PRT
115 <213> ORGANISM: Dahlia merckii
117 <400> SEQUENCE: 5
118 Lys Lys Ala Glu Lys Leu Ala Gln Asp Lys Leu Lys Ala Glu Gln Leu
119  1      5      10      15
121 Ala Gln Asp Lys Leu Asn Ala Gln Lys Leu Asp Arg Asp Ala Lys Lys
122          20      25      30
124 Val Val Pro Asn Val Glu His Pro
125          35      40
129 <210> SEQ ID NO: 6
130 <211> LENGTH: 44
131 <212> TYPE: PRT
132 <213> ORGANISM: Artificial Sequence
134 <220> FEATURE:
135 <223> OTHER INFORMATION: Description of Artificial Sequence: Linker
136     propeptide
138 <400> SEQUENCE: 6
139 Lys Lys Ala Glu Lys Leu Ala Gln Asp Lys Leu Lys Ala Glu Gln Leu
140  1      5      10      15
142 Ala Gln Asp Lys Leu Asn Ala Gln Lys Leu Asp Arg Asp Ala Lys Lys

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/763,076
 DATE: 03/02/2001
 TIME: 13:01:10

Input Set : A:\PPD50348 US SEQ LIST.txt
 Output Set: N:\CRF3\03022001\I763076.raw

```

143          20          25          30
145 Val Val Pro Asn Val Glu His Pro Ile Gly Lys Arg
146          35          40
150 <210> SEQ ID NO: 7
151 <211> LENGTH: 20
152 <212> TYPE: PRT
153 <213> ORGANISM: Artificial Sequence
155 <220> FEATURE:
156 <223> OTHER INFORMATION: Description of Artificial Sequence: Linker
157     propeptide
159 <400> SEQUENCE: 7
160 Ala Ser Thr Thr Val Asp His Gln Ala Asp Val Ala Ala Thr Lys Thr
161   1          5          10          15
163 Ile Gly Lys Arg
164          20
168 <210> SEQ ID NO: 8
169 <211> LENGTH: 31
170 <212> TYPE: PRT
171 <213> ORGANISM: Amaranthus caudatus
174 <400> SEQUENCE: 8
175 Ala Ser Thr Thr Val Asp His Gln Ala Asp Val Ala Ala Thr Lys Thr
176   1          5          10          15
178 Ala Lys Asn Pro Thr Asp Ala Lys Leu Ala Gly Ala Gly Ser Pro
179          20          25          30
183 <210> SEQ ID NO: 9
184 <211> LENGTH: 522
185 <212> TYPE: DNA
186 <213> ORGANISM: Artificial Sequence
188 <220> FEATURE:
189 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
190     sequence
192 <220> FEATURE:
193 <221> NAME/KEY: CDS
194 <222> LOCATION: (76)..(513)
196 <400> SEQUENCE: 9
197 ctcgagtatt ttacaacaa ttaccaacaa caacaacaa caaacaacat tacaattact 60
199 atttacaatt acacc atg gtg aat cgg tgg gtt gcg ttc tcc gcg ttc gtt 111
200           Met Val Asn Arg Ser Val Ala Phe Ser Ala Phe Val
201           1          5          10
203 ctg atc ctt ttc gtg ctc gcc atc tca gat atc gca tcc gtt agt gga 159
204 Leu Ile Leu Phe Val Leu Ala Ile Ser Asp Ile Ala Ser Val Ser Gly
205           15          20          25
207 gaa cta tgc gag aaa gct agc aag acg tgg tgg ggc aac tgt ggc aac 207
208 Glu Leu Cys Glu Lys Ala Ser Lys Thr Trp Ser Gly Asn Cys Gly Asn
209           30          35          40
211 acg gga cat tgt gac aac caa tgt aaa tca tgg gag ggt gcg gcc cat 255
212 Thr Gly His Cys Asp Asn Gln Cys Lys Ser Trp Glu Gly Ala Ala His
213 45          50          55          60
215 gga gcg tgt cat gtg cgt aac ggg aaa cac atg tgt ttc tgt tac ttc 303

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/763,076

DATE: 03/02/2001

TIME: 13:01:10

Input Set : A:\PPD50348 US SEQ LIST.txt

Output Set: N:\CRF3\03022001\I763076.raw

```

216 Gly Ala Cys His Val Arg Asn Gly Lys His Met Cys Phe Cys Tyr Phe
217          65          70          75
219 aat tgt tcc aac gct gct gac gag gtg gct acc cca gag gac gtg gag 351
220 Asn Cys Ser Asn Ala Ala Asp Glu Val Ala Thr Pro Glu Asp Val Glu
221          80          85          90
223 cca gga cag aag ttg tgc caa agg cca agt ggg aca tgg tca gga gtc 399
224 Pro Gly Gln Lys Leu Cys Gln Arg Pro Ser Gly Thr Trp Ser Gly Val
225          95          100          105
227 tgt gga aac aat aac gca tgc aag aat cag tgc att aga ctt gag aaa 447
228 Cys Gly Asn Asn Asn Ala Cys Lys Asn Gln Cys Ile Arg Leu Glu Lys
229          110          115          120
231 gca cga cat gga tct tgc aac tat gtc ttc cca gct cac aag tgt atc 495
232 Ala Arg His Gly Ser Cys Asn Tyr Val Phe Pro Ala His Lys Cys Ile
233 125          130          135          140
235 tgc tac ttt cct tgt taa taggagctc 522
236 Cys Tyr Phe Pro Cys
237          145
240 <210> SEQ ID NO: 10
241 <211> LENGTH: 145
242 <212> TYPE: PRT
243 <213> ORGANISM: Artificial Sequence
245 <220> FEATURE:
246 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
247 sequence
249 <400> SEQUENCE: 10
250 Met Val Asn Arg Ser Val Ala Phe Ser Ala Phe Val Leu Ile Leu Phe
251 1          5          10          15
253 Val Leu Ala Ile Ser Asp Ile Ala Ser Val Ser Gly Glu Leu Cys Glu
254          20          25          30
256 Lys Ala Ser Lys Thr Trp Ser Gly Asn Cys Gly Asn Thr Gly His Cys
257          35          40          45
259 Asp Asn Gln Cys Lys Ser Trp Glu Gly Ala Ala His Gly Ala Cys His
260          50          55          60
262 Val Arg Asn Gly Lys His Met Cys Phe Cys Tyr Phe Asn Cys Ser Asn
263 65          70          75          80
265 Ala Ala Asp Glu Val Ala Thr Pro Glu Asp Val Glu Pro Gly Gln Lys
266          85          90          95
268 Leu Cys Gln Arg Pro Ser Gly Thr Trp Ser Gly Val Cys Gly Asn Asn
269          100          105          110
271 Asn Ala Cys Lys Asn Gln Cys Ile Arg Leu Glu Lys Ala Arg His Gly
272          115          120          125
274 Ser Cys Asn Tyr Val Phe Pro Ala His Lys Cys Ile Cys Tyr Phe Pro
275          130          135          140
277 Cys
278 145
282 <210> SEQ ID NO: 11
283 <211> LENGTH: 534
284 <212> TYPE: DNA
285 <213> ORGANISM: Artificial Sequence

```

RAW SEQUENCE LISTING DATE: 03/02/2001
 PATENT APPLICATION: US/09/763,076 TIME: 13:01:10

Input Set : A:\PPD50348 US SEQ LIST.txt
 Output Set: N:\CRF3\03022001\I763076.raw

```

287 <220> FEATURE:
288 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
289     sequence
291 <220> FEATURE:
292 <221> NAME/KEY: CDS
293 <222> LOCATION: (76)..(525)
295 <400> SEQUENCE: 11
296 ctcgagtatt tttaacaacaa ttaccaacaa caacaaacaa caaacaacat tacaattact 60
298 atttacaatt acacc atg gtg aat cgg tcg gtt gcg ttc tcc gcg ttc gtt 111
299     Met Val Asn Arg Ser Val Ala Phe Ser Ala Phe Val
300           1             5             10
302 ctg atc ctt ttc gtg ctc gcc atc tca gat atc gca tcc gtt agt gga 159
303 Leu Ile Leu Phe Val Leu Ala Ile Ser Asp Ile Ala Ser Val Ser Gly
304           15             20             25
306 gaa cta tgc gag aaa gct agc aag acg tgg tcg ggc aac tgt ggc aac 207
307 Glu Leu Cys Glu Lys Ala Ser Lys Thr Trp Ser Gly Asn Cys Gly Asn
308           30             35             40
310 acg gga cat tgt gac aac caa tgt aaa tca tgg gag ggt gcg gcc cat 255
311 Thr Gly His Cys Asp Asn Gln Cys Lys Ser Trp Glu Gly Ala Ala His
312 45           50             55             60
314 gga gcg tgt cat gtg cgt aac ggg aaa cac atg tgt ttc tgt tac ttc 303
315 Gly Ala Cys His Val Arg Asn Gly Lys His Met Cys Phe Cys Tyr Phe
316           65             70             75
318 aat tgt aaa aaa gcc gaa aag ctt gct caa gac aaa ctt aaa gcc gaa 351
319 Asn Cys Lys Lys Ala Glu Lys Leu Ala Gln Asp Lys Leu Lys Ala Glu
320           80             85             90
322 caa ctc atc gga aag agg cag aag ttg tgc caa agg cca agt ggg aca 399
323 Gln Leu Ile Gly Lys Arg Gln Lys Leu Cys Gln Arg Pro Ser Gly Thr
324           95             100            105
326 tgg tca gga gtc tgt gga aac aat aac gca tgc aag aat cag tgc att 447
327 Trp Ser Gly Val Cys Gly Asn Asn Asn Ala Cys Lys Asn Gln Cys Ile
328           110            115            120
330 aga ctt gag aaa gca cga cat gga tct tgc aac tat gtc ttc cca gct 495
331 Arg Leu Glu Lys Ala Arg His Gly Ser Cys Asn Tyr Val Phe Pro Ala
332 125           130            135            140
334 cac aag tgt atc tgc tac ttt cct tgt taa taggagctc 534
335 His Lys Cys Ile Cys Tyr Phe Pro Cys
336           145
339 <210> SEQ ID NO: 12
340 <211> LENGTH: 149
341 <212> TYPE: PRT
342 <213> ORGANISM: Artificial Sequence
344 <220> FEATURE:
345 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
346     sequence
348 <400> SEQUENCE: 12
349 Met Val Asn Arg Ser Val Ala Phe Ser Ala Phe Val Leu Ile Leu Phe
350 1             5             10             15
352 Val Leu Ala Ile Ser Asp Ile Ala Ser Val Ser Gly Glu Leu Cys Glu

```

FJI

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/763,076

DATE: 03/02/2001

TIME: 13:01:11

Input Set : A:\PPD50348 US SEQ LIST.txt

Output Set: N:\CRF3\03022001\I763076.raw

L:14 M:270 C: Current Application Number differs, Replaced Application Number
L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:415 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:2060 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52
L:2252 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:66